المحور الثاني: التوزيعات الاحتمالية المستمرة

I. المتغير العشوائي المتصل وتوزيع الاحتمالي:

❖ المتغير العشوائي المتصل

مثلما رأينا سابقا، يقال على المتغير العشوائي X غير المنفصل أنه متغير عشوائي مستمر، وهو المتغير الذي يكون المجال المقابل له غير قابل للعد، أي أن قيمه تأخذ مجال من الأعداد الحقيقية بين قيمتين محددتين (b و a)، أو غير محددتين (أي بين∞- و ∞+)، وتكون دالة التوزيع خاصته على الشكل التالي:

$$F(x) = f(X \leq x) = \int_{-\infty}^{x} f(u)du \ldots\ldots\ldots\ldots\ldots (1)$$

حيث الدالة f(x) لها الخصائص التالية:

1. $f(x) \geq 0 \qquad\qquad \forall X_i \quad a \leq X_i \leq b$
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$

❖ التوزيع الاحتمالي المستمر:

من التعريف السابق نجد أن المتغير العشوائي المتصل x له احتمال صفري 0 لأن يكون مساويا لقيمة معينة. بينما فترة الاحتمال للمتغير x أن يقع بين قيمتين مختلفتين a,b ويكون:

$$P(a \leq x \leq b) = \int_{a}^{b} f(x)dx \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

والدالة التي تحقق هذه المتطلبات f(x) تسمى بدالة الاحتمال أو دالة الكثافة الاحتمالية للمتغير x.

ولإيجاد الاحتمال الواقع ضمن المجال (a, b) نقوم بحساب المساحة تحت المنحنى أي: P(a≤x≤b)

مثال 01:

لتكن الدالة: $f(X_i) = \begin{cases} 3X^2 & 0 \leq X \leq 1 \\ 0 & \frac{o}{w} \end{cases}$

- اثبت أن الدالة $f(X_i)$ تمثل دالة كثافة احتمالية.
- احسب قيمة الاحتمال $P\left(0 \leq x \leq \frac{1}{2}\right)$

الحل:

لإثبات أن الدالة $f(X_i)$ هي دالة كثافة احتمالية، ينبغي أن تحقق الخاصيتين:

1. $f(x) \geq 0 \qquad \forall X_i \quad 0 \leq X_i \leq 1$
2. $\int_0^1 f(x)dx = \int_0^1 3X^2 dx = X^3|_0^1 = 1^3 - 0^3 = 1$

إذن الدالة $f(X_i)$ تمثل دالة كثافة احتمالية.

- حساب قيمة الاحتمال $P\left(0 \leq x \leq \frac{1}{2}\right)$

$$P\left(0 \leq x \leq \frac{1}{2}\right) = \int_0^{\frac{1}{2}} 3X^2 dx = X^3|_0^{\frac{1}{2}} = \frac{1}{8}$$

❖ **دوال التوزيع للمتغيرات العشوائية المتصلة:**

نتذكر تعريف دالة التوزيع التراكمية أو دالة التوزيع للمتغير العشوائي x وهي:

$$F(x) = f(X \leq x) \ldots\ldots\ldots\ldots\ldots (3)$$

حيث أن x أي رقم حقيقي، أي أن $(-\infty \leq x \leq +\infty)$ وبالتالي:

$$F(x) = \int_{-\infty}^{x} f(x)dx \ldots\ldots\ldots\ldots\ldots (4)$$

وتوجد علاقة بين دالة التوزيع F(x) ودالة الكثافة الاحتمالية f(x)، حيث أن اشتقاق دالة التوزيع يعطينا دالة الكثافة الاحتمالية.

❖ **القيم المتوقعة (التوقع)**

التوقع للمتغير العشوائي المتصل x الذي له دالة كثافة احتمالية f(x) يكون على الشكل:

$$E(x) = \mu = \int_{-\infty}^{+\infty} x * f(x)dx \ldots\ldots\ldots\ldots\ldots (5)$$

وهو مقدار ثابت لأن التكامل محدود ويعتمد على الثوابت التي تتضمنها دالة الكثافة.

❖ **التباين والانحراف المعياري**

إذا كان x متغيرا عشوائيا متصلا له دالة كثافة احتمالية f(x) فإن التباين يكون على الشكل:

$$\sigma_x^2 = E[(x-\mu)^2] = \int_{-\infty}^{+\infty} (x)^2 f(x)dx - \mu^2 \dots\dots\dots\dots\dots (6)$$

ودائما الانحراف المعياري هو جذر التباين.

مثال 02:

لتكن دالة الكثافة الاحتمالية:
$$f(X_i) = \begin{cases} 3X^2 & 0 \le X \le 1 \\ 0 & \frac{o}{w} \end{cases}$$

- احسب كل من التوقع الرياضي والتباين.

الحل:

- التوقع الرياضي:

$$E(x) = \mu = \int_0^1 x * f(x)dx = \int_0^1 X * 3X^2 dx = \int_0^1 3X^3 dx = \frac{3X^3}{4}\Big|_0^1 = \frac{3}{4}$$

- التباين:

$$\sigma_x^2 = \int_{-\infty}^{+\infty} (x)^2 f(x)dx - \mu^2 = \int_0^1 3X^4 dx - \mu^2 = \frac{3X^5}{5}\Big|_0^1 - \left(\frac{3}{4}\right)^2 = \frac{3}{80}$$

II. التوزيعات الاحتمالية الأكثر استخداما في التطبيقات الاحصائية.

1) التوزيع الطبيعي:

يعود الفضل في اكتشاف هذا التوزيع إلى العالم الرياضي الانجليزي دي مويفر (De-Moivre) عام 1733، وكان اول من استخدم التوزيع الطبيعي في دراسة الاخطاء المحتملة في القياس كل من العالمين الرياضيين لابلاس (Laplace) وغوص (Gauss) علم 1809.

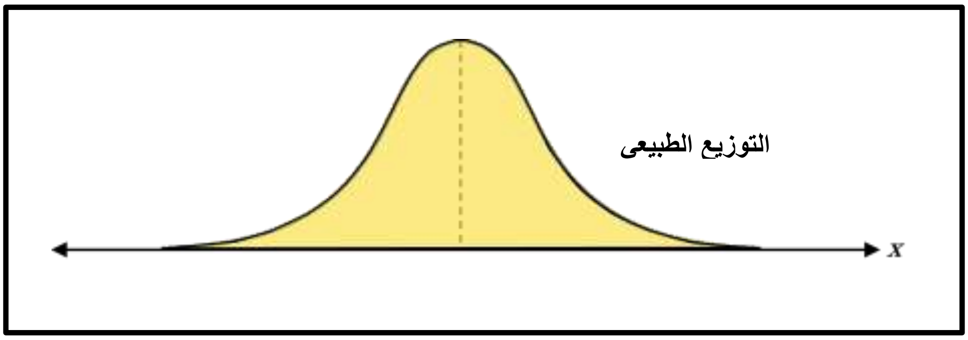
يعتبر التوزيع الطبيعي من أهم التوزيعات المتصلة حيث يلعب دوراً أساسياً في عملية المعاينة Sampling، كما يستخدم لوصف الأنماط التكرارية للعديد من الظواهر الإحصائية مثل التغيرات الطبيعية التي تحدث للإنسان والحيوان والعوامل البيئية بشكل عام. والتوزيع الطبيعي يدرس سلوك المتغيرات العشوائية المتصلة مثل درجة الحرارة والطول والوزن والدخل والأخطاء العشوائية الناتجة عند تحليل الانحدار.

ولرسم التوزيع الطبيعي نستخدم المنحنى الاحتمالي للتوزيع الطبيعي، فإذا كانت x متغيراً عشوائياً متصلاً يتبع توزيعاً طبيعياً متوسطه μ و تباينه σ^2 يمكن رسم المنحنى الاحتمالي للتوزيع الطبيعي باستخدام المعادلة:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad ; \; -\infty \leq x \leq \infty$$

ويقال إن x تتبع توزيعاً طبيعياً متوسطه μ و تباينه $\sigma2$. و تكتب اختصارا:

$$X \sim N(\mu , \sigma^2)$$



وحيث أن التوزيع الطبيعي يعبر عن متغيرات عشوائية متصلة فإننا لا نستطيع استخدام دالة التوزيع له لإيجاد احتمال أن يأخذ المتغير العشوائي الذى يتبعه لقيمة معينة، حيث أن هذا الاحتمال يساوى صفراً . ولكن يمكن استخدام دالة التوزيع الطبيعي لإيجاد احتمال أن يأخذ المتغير العشوائي أقل من قيمة معينة أو أكثر من قيمة معينة أو تقع قيمته بين قيمتين معلومتين.

وترجع أهمية التوزيع الطبيعي إلى الأسباب الآتية :

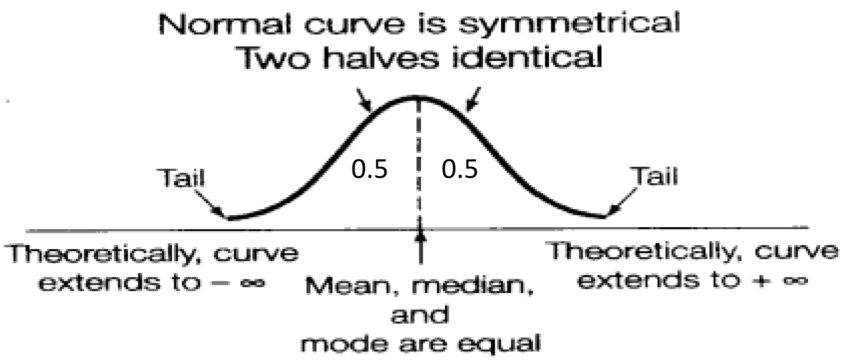
- يستخدم التوزيع الطبيعي في وصف العديد من المتغيرات العشوائية في الواقع العملي مثل الأطوال والأوزان لمجموعة من الأشخاص ، درجات الطلاب في امتحان مقرر معين ، أخطاء القياس الناتجة من تجربة ما ، حيث تتوزع القياسات المشاهدة (التكرارات) بشكل متماثل حول قيمة مركزية هي قيمة متوسط التوزيع وتقل هذه القياسات تدريجيا كلما ابتعدنا عن هذه القيمة إلى جهة اليمين أو إلى جهة اليسار ، حيث يأخذ منحنى التوزيع الشكل الجرسي.
- يعتبر التوزيع الطبيعي تقريبا مفيدا للعديد من التوزيعات المتقطعة وغير المتقطعة مثل توزيع ذي الحدين وتوزيع بواسون ....

- يعتبر التوزيع الطبيعي حجر الزاوية في الاستدلال الإحصائي حيث تتوزع معالم المجتمع المقدرة من العينة مثل المتوسط مثلا كالتوزيع الطبيعي وهذا ما سوف تناوله في معرض حديثنا عن نظرية النهاية المركزية.

❖ خصائص التوزيع الطبيعي:

يتميز التوزيع الطبيعي بالخصائص الآتية:

1. منحنى التوزيع الطبيعي منحنى متماثل حول متوسط التوزيع والذى يرمز له بالرمز μ .
2. المتوسط والوسيط والمنوال للتوزيع الطبيعي متساوية القيمة.
3. المساحة تحت منحنى التوزيع الطبيعي تساوى الوحدة المربعة وهذه المساحة تعبر عن مجموع الاحتمالات . وكما ذكرنا أن الخط الرأسي الساقط من قمة المنحنى على المحور الأفقي عند المتوسط يقسم المساحة إلى نصفين متساويين 50% من المساحة الكلية على يمين الخط العمودي و% 50 من المساحة الكلية على يساره.



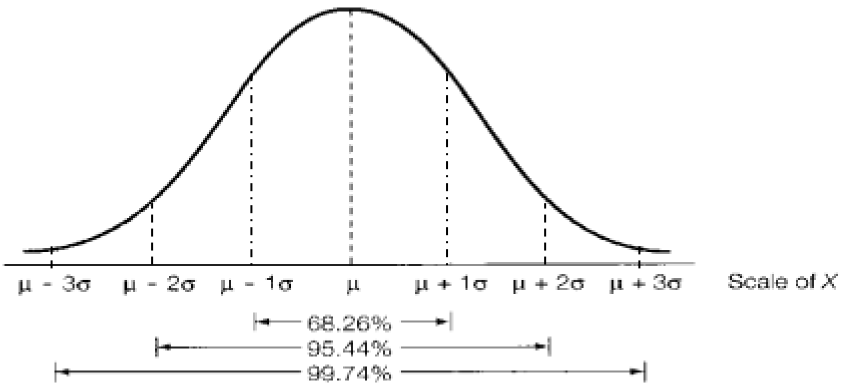
4. إذا أسقطنا عمودين على بعد انحراف معياري واحد إلى يمين ويسار متوسط التوزيع فإن المساحة التي يحصرها هذين العمودين تمثل 68.26% تقريبا من المساحة الكلية تحت المنحنى. كذلك نجد أن المساحة التي يحصرها العمودان المقامان عند 2 انحراف معياري على جانبي متوسط التوزيع تمثل 95.44% تقريبا من المساحة الكلية تحت المنحنى. أما المساحة التي يحصرها العمودان المقامان عند 3 انحرافات معيارية على جانبي متوسط التوزيع فتمثل 99.74% تقريبا من المساحة الكلية تحت المنحنى.

هذا ويمكن التعبير عما سبق رمزيا كالآتي:

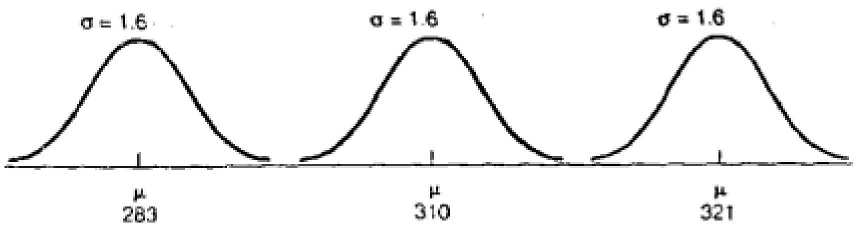
$$p\ (\ \mu - 1\sigma < x < \mu + 1\ \sigma\ ) = 0.6826$$

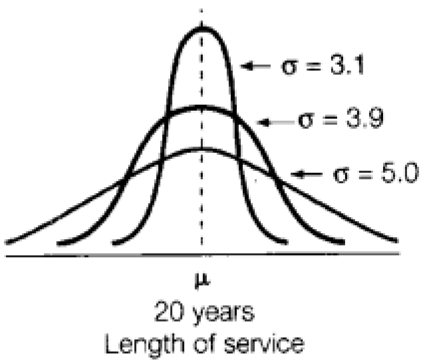
$$p\ (\ \mu - 2\sigma < x < \mu + 2\ \sigma\ ) = 0.9544$$

$$p\ (\ \mu - 3\sigma < x < \mu + 3\ \sigma\ ) = 0.99.74$$



5. التوزيع الطبيعي ليس توزيعاً وحيدا ولكنه يمثل عائلة من التوزيعات الطبيعية، وكل توزيع من هذه العائلة يمكن تحديده وتعريفه بدقة متى عرفت معالمه وهى المتوسط ويرمز له بالرمز μ، والانحراف المعياري ويرمز له بالرمز σ، حيث تحدد قيمتي هتين المعلمتين شكل منحنى التوزيع، حيث تحدد قيمة المتوسط μ موقع التوزيع على المحور الأفقي من نقطة الأصل، كما تحدد قيمة الانحراف المعياري مدى تشتت التوزيع فكلما كانت σ كبيرة كلما زاد تشتت القيم وبالتالي اتساع المنحنى.





❖ التوزيع الطبيعي المعياري (كحالة خاصة):

التوزيع الطبيعي المعياري هو التوزيع الطبيعي بمتوسط معدوم $\mu = 0$ وتباين $\sigma^2 = 1$. أي أن أي قيمة من قيم المتغير X يمكن تحويلها إلى المتغير Z الذي يتبع التوزيع الطبيعي المعياري باستخدام الصيغة التالية:

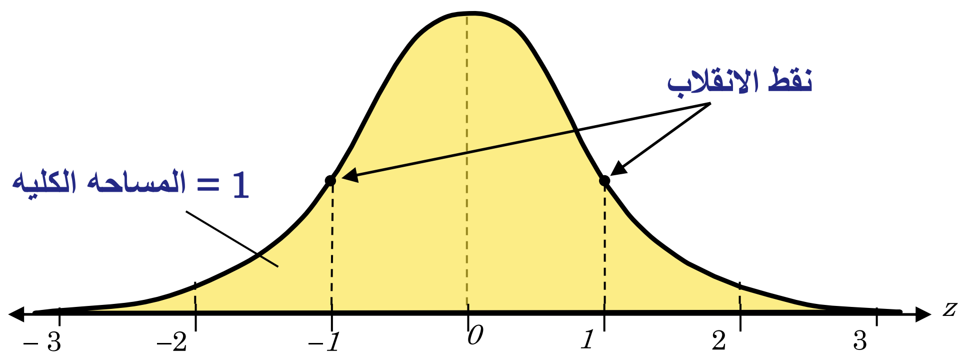
$$Z = \frac{X - \mu}{\sigma}$$

ويمكن رسم المنحنى الاحتمالي للتوزيع الطبيعي المعياري باستخدام دالة الكثافة الاحتمالية التالية:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad -\infty \le z \le \infty.$$

والشكل الآتي يوضح المنحنى الاحتمالي للتوزيع الطبيعي المعياري:

المنحنى الإحتمالى للتوزيع الطبيعى المعياري



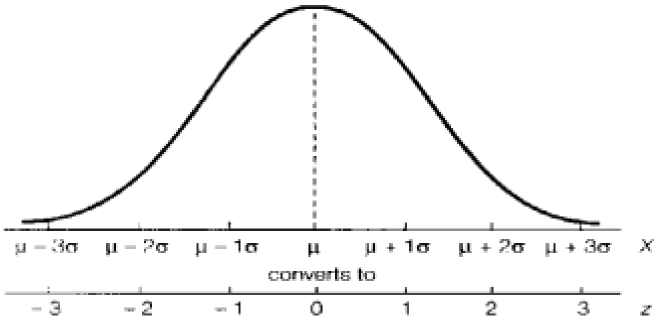
❖ خصائص منحنى التوزيع الطبيعي المعياري:

1. المساحة الكلية تحت منحنى التوزيع الطبيعي المعياري تساوى الواحد الصحيح بحيث يقسم الخط العمودي الواصل من قمة المنحنى إلى الصفر (متوسط التوزيع) المساحة إلى قسمين متساويين مساحة كل منهما تساوى 0.50.
2. منحنى التوزيع الطبيعي المعياري متماثل حول متوسطه، وبالتالي فإن التواءه يساوى صفر وتفرطحه يساوى 3.
3. المساحة المحصورة بين ±1 درجة معيارية تساوى 68.26% تقريبا من المساحة الكلية تحت المنحنى. والمساحة المحصورة بين 2 ± درجة معيارية تساوي 95.44 % تقريبا من المساحة الكلية تحت المنحنى. أما المساحة المحصورة بين 3 ± درجات معيارية تساوي 99.74 % تقريبا من المساحة الكلية تحت المنحنى. هذا ويمكن التعبير عن ذلك رمزيا كالآتي:

$$P(-1 < Z < 1) = 0.6826$$

P (-2 < Z < 2) = 0.9544

P ((-3 < Z < 3) = 0.9974



أمثلة:

أولا: حساب الاحتمال في حالة التوزيع الطبيعي المعياري

إذا كان Z متغير عشوائيا متصلا يتبع التوزيع الطبيعي المعياري فأوجد:

1. $P(Z \leq 1.54)$
2. $P(-1.8 \leq Z \leq 0)$
3. $P(1 \leq Z \leq 2)$

الحل:

1. $P(Z < 1.54) = 0.9382$

من الجدول الإحصائي للتوزيع الطبيعي المعياري نستخرج القيمة:

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505

2. $P(- 1.8 \leq Z \leq 0)$

$= P(Z \leq 0) - P(Z \leq -1.8)$

$= 0.5 - 0.359$

$= 0.4641$

z	0.00	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985

z	0.00	0.01
−3.4	0.0003	0.0003
−3.3	0.0005	0.0005
−3.2	0.0007	0.0007
−3.1	0.0010	0.0009
−3.0	0.0013	0.0013
−2.9	0.0019	0.0018
−2.8	0.0026	0.0025
−2.7	0.0035	0.0034
−2.6	0.0047	0.0045
−2.5	0.0062	0.0060
−2.4	0.0082	0.0080
−2.3	0.0107	0.0104
−2.2	0.0139	0.0136
−2.1	0.0179	0.0174
−2.0	0.0228	0.0222
−1.9	0.0287	0.0281
−1.8	0.0359	0.0351
−1.7	0.0446	0.0436

.3 $P(1 \leq Z \leq 2)$

$= P(Z < 2) - P(Z < 1)$

$P(1 < Z < 2) =$

$= P(Z < 2) - P(Z < 1)$

$= 0.1359$

z	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700
0.5	0.6915	0.6950	0.6985	0.7019	0.7054
0.6	0.7257	0.7291	0.7324	0.7357	0.7389
0.7	0.7580	0.7611	0.7642	0.7673	0.7704
0.8	0.7881	0.7910	0.7939	0.7967	0.7995
0.9	0.8159	0.8186	0.8212	0.8238	0.8264
1.0	0.8413	0.8438	0.8461	0.8485	0.8508
1.1	0.8643	0.8665	0.8686	0.8708	0.8729
1.2	0.8849	0.8869	0.8888	0.8907	0.8925
1.3	0.9032	0.9049	0.9066	0.9082	0.9099
1.4	0.9192	0.9207	0.9222	0.9236	0.9251
1.5	0.9332	0.9345	0.9357	0.9370	0.9382
1.6	0.9452	0.9463	0.9474	0.9484	0.9495
1.7	0.9554	0.9564	0.9573	0.9582	0.9591
1.8	0.9641	0.9649	0.9656	0.9664	0.9671
1.9	0.9713	0.9719	0.9726	0.9732	0.9738
2.0	0.9772	0.9778	0.9783	0.9788	0.9793
2.1	0.9821	0.9826	0.9830	0.9834	0.9838
2.2	0.9861	0.9864	0.9868	0.9871	0.9875
2.3	0.9893	0.9896	0.9898	0.9901	0.9904
2.4	0.9918	0.9920	0.9922	0.9925	0.9927

ثانيا: حساب الاحتمال في حالة التوزيع الطبيعي العادي

مثال 03:

إذا كان العمر الافتراضي للمصابيح المنتجة في شركة محلية يتبع التوزيع الطبيعي بمتوسط 1000 ساعة وانحراف معياري قدره 100 ساعة، فإذا تم اختيار مصباح من انتاج هذه الشركة بطريقة عشوائية ، فأوجد احتمال أن يكون عمرها الافتراضي:

1. بين 1000 ، 1150 ساعة.
2. أقل من 930 ساعة.
3. أكبر من 780 ساعة.
4. بين 700 ، 1200 ساعة.
5. بين 750 ، 850 ساعة.

الحل:

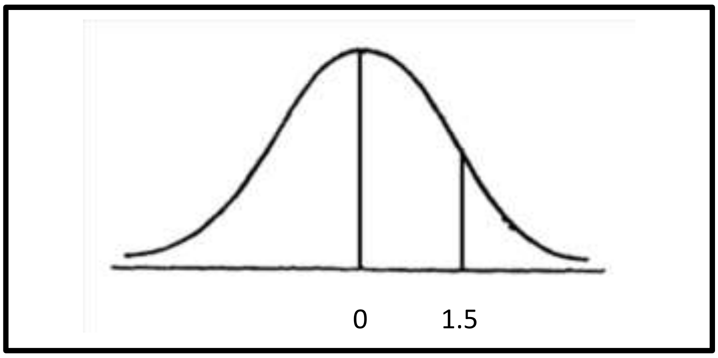
نفرض أن المتغير العشوائي x يعبر عن العمر الافتراضي للمصابيح المنتجة، أي أن:

المتوسط μ = 1000 ساعة

الانحراف المعياري σ = 100 ساعة

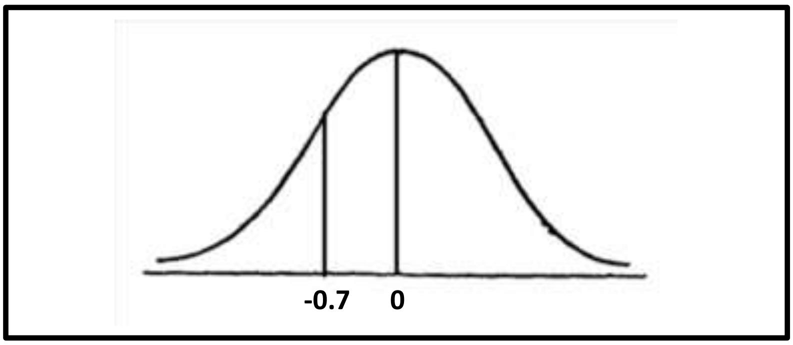
1. احتمال أن يكون العمر الافتراضي للمصباح بين 1000 ، 1150 ساعة

$$P(1000 \le X \le 1150) = P\left(\frac{1000-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{1150-1000}{100}\right)$$
$$= P(0 \le Z \le 1.5) = 0.4332$$



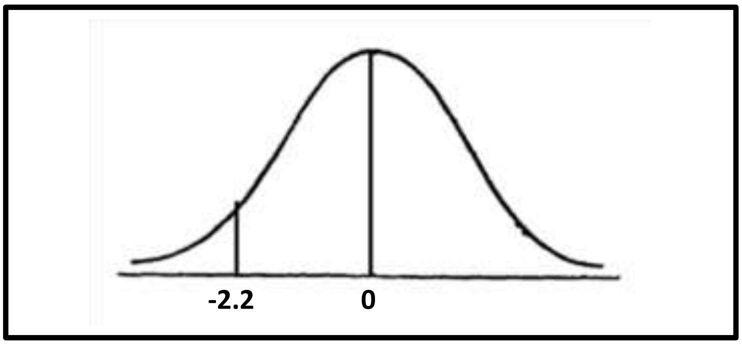
1. احتمال أن يكون العمر الافتراضي للمصباح أقل من 930 ساعة

$$P(X < 930) = P\left(\frac{X-\mu}{\sigma} < \frac{930-1000}{100}\right) = p\left(Z < \frac{-70}{100}\right)$$
$$= p(Z < -0.7) = 0.5 - 0.2580 = 0.2420$$



3. احتمال أن يكون العمر الافتراضي للمصباح أكبر من 780 ساعة

$$P(X > 780) = P\left(\frac{X-\mu}{\sigma} > \frac{780-1000}{100}\right) = P\left(Z > \frac{-220}{100}\right)$$
$$= P(Z > -2.2) = 0.5 + P(-2.2 < Z < 0)$$
$$= 0.5 + P(0 < Z < 2.2)$$
$$= 0.5 + 0.4861 = 0.9861$$



4. احتمـــال أن يكـــون العمـــر الافتراضـــي للمصـــباح بين 700 ، 1200 ساعة

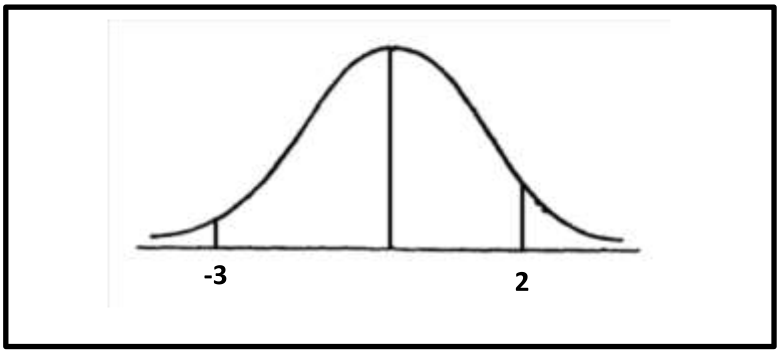
$$p(700 \le X \le 1200) = P\left(\frac{700-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{1200-1000}{100}\right)$$

$$= P\left(\frac{-300}{100} \le Z \le \frac{200}{100}\right) = P(-3 \le Z \le 2)$$

$$= P(-3 \le Z \le 0) + P(0 \le Z \le 2)$$

$$= P(0 \le Z \le 3) + P(0 \le Z \le 2)$$

$$= 0.4987 + 0.4772 = 0.9759$$



5. احتمـــال أن يكـــون العمـــر الافتراضـــي للمصـــباح بين 750 ، 850 ساعة

$$P(750 \le X \le 850) = P\left(\frac{750-1000}{100} \le \frac{X-\mu}{\sigma} \le \frac{850-1000}{100}\right)$$

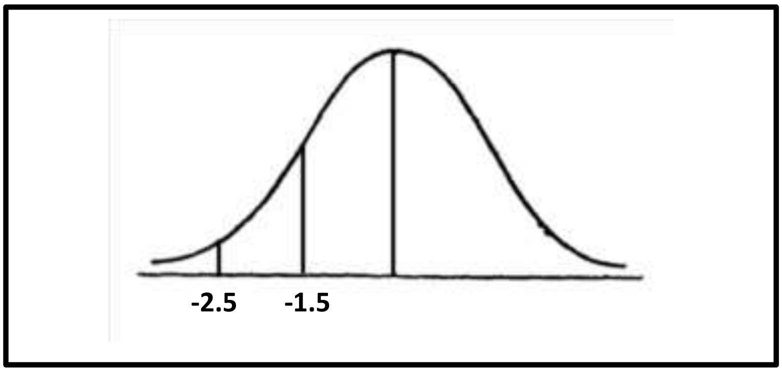
$$= P\left(\frac{-250}{100} \le Z \le \frac{-150}{100}\right)$$

$$= P(-2.50 \le Z \le -1.50)$$

$$= P(-2.50 \le Z \le 0) - P(-1.50 \le Z \le 0)$$

$$= P(0 \le Z \le 2.50) - P(0 \le Z \le 1.50)$$

$$= 0.4938 - 0.4332 = 0.0606$$



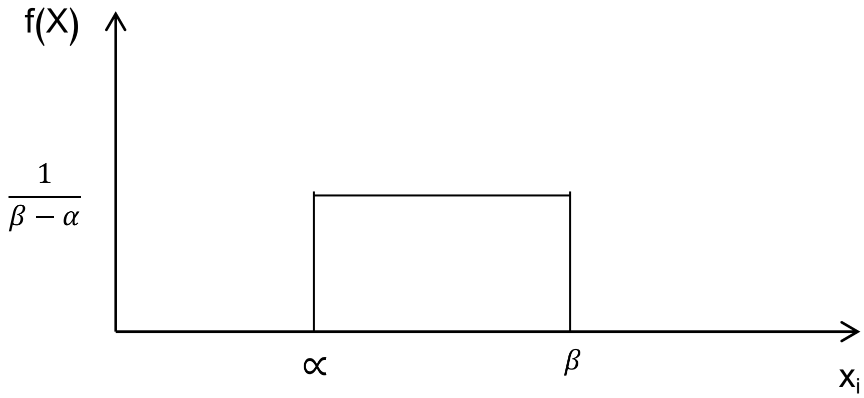
2) التوزيع المنتظم Uniform Distribution:

يعد التوزيع المنتظم من التوزيعات الاحتمالية المتصلة المهمة جدا لكثير من التطبيقات في الواقع العملي، إذ يستخدم لحساب الاحتمالات المناسبة لهذه التطبيقات، مثل دراسة احتمال وصول البواخر إلى الموانئ لتفريغ الحمولة، أو وصول الشاحانات إلى محطات التفريغ ...الخ.

وبافتراض ان لدينا متغير عشوائي x متصل، يتوزع وفق التوزيع المنتظم بالمعلمتين (α, β)، فإن دالة التوزيع الاحتمالي f(x) للمتغير العشوائي x تأخذ الشكل التالي:

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 0 & o/w \end{cases}$$

والشكل التالي يوضح دالة الكثافة الاحتمالية للتوزيع المنتظم:



وغالبا ما يعبر عن التوزيع المنتظم اختصارا بالاصطلاح الاتي: $x \sim U(\alpha, \beta)$

دالة التوزيع التراكمي للتوزيع المنتظم:

$$F(x) = P(X \leq x) = \int_{\alpha}^{x} f(x)dx$$

$$= \int_{\alpha}^{x} \frac{1}{\beta - \alpha} dx$$

$$= \frac{1}{\beta - \alpha} \int_{\alpha}^{x} dx$$

$$= \frac{1}{\beta - \alpha} X|_{\alpha}^{x}$$

$$F(x) = \frac{x - \alpha}{\beta - \alpha}$$

وعليه تكتب دالة التوزيع التجميعية في شكلها النهائي على النحو التالي:

$$F(x) = P(X \leq x) = \begin{cases} 0 & x \leq \alpha \\ \dfrac{x - \alpha}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 1 & x \geq \beta \end{cases}$$

❖ خصائص التوزيع المنتظم:

- الوسط الحسابي: $\mu_x = \frac{\alpha+\beta}{2}$
- التباين: $\sigma_x^2 = \frac{(\beta-\alpha)^2}{12}$
- الانحراف المعياري: $\sigma_x = \frac{\beta-\alpha}{\sqrt{12}}$

مثال 04:

إذا كان وقت الوصول الحقيقي لباخرة ما إلى ميناء الجزائر، هو متغير عشوائي مستمر يتبع التوزيع المنتظم ذو المعلمتين $\beta = 20 \quad \alpha = 0$، أي أن $x \sim U(0,20)$

- احسب دالة الكثافة الاحتمالية ودالة التوزيع التجميعية للمتغير العشوائي x.
- ما هو احتمال وصول الباخرة خلال خمس 5 دقائق الأخيرة على الاقل.
- ما هو احتمال وصول الباخرة خلال ال10 دقائق الاخيرة؟
- احسب القيمة المتوقعة والتباين لوقت وصول الباخرة.

الحل:

دالة الكثافة الاحتمالية:

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & \alpha \leq x \leq \beta \\ 0 & o/w \end{cases}$$

$$f(x) = \begin{cases} \frac{1}{20} & 0 \le x \le 20 \\ 0 & o/w \end{cases}$$

دالة التوزيع التجميعية:

$$F(x) = \begin{cases} 0 & x \le 0 \\ \frac{x}{20} & 0 \le x \le 20 \\ 1 & x \ge 20 \end{cases}$$

حساب الاحتمال:

$$P(X > 15) = 1 - P(X \le 15)$$

$$P(X > 15) = 1 - F(15)$$

$$= 1 - \frac{15}{20}$$

$$P(X > 15) = 0.25$$

$$P(10 \le X \le 20) = P(X \le 20) - P(X \le 10)$$

$$= F(20) - F(10)$$

$$= 1 - 0.5 = 0.5$$

الوسط الحسابي والتباين:

$$\mu_x = \frac{\alpha + \beta}{2} = \frac{0 + 20}{2} = 10 \text{ دقائق}$$

$$\sigma_x^2 = \frac{(\beta - \alpha)^2}{12} = \frac{(20 - 0)^2}{12} = 33.3$$

3) توزيع قاما Gamma Distribution:

يعد توزيع قاما من التوزيعات الشائعة الاستخدام في بعض التطبيقات الاحصائية، مثل داالة البقاء، ويعالج عادة المتغيرات العشوائية التي تكون قيمها موجبة دائما، مثل فترة بقاء انسان مصاب بداء عضال على قيد الحياة، الفترة الزمنية لفحص مريض في إحدى العيادات، الفترة الزمنية بين وصول قطارين متتاليين إلى إحدى المحطات...الخ.

وبافتراض أن لدينا متغير عشوائي متصل x يتوزع وفق توزيع قاما، فإن دالة توزيعه الاحتمالي تأخذ الشكل:

$$f(x) = \begin{cases} \frac{1}{\beta^{\propto} \Gamma \propto} x^{\propto-1} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & o/w \end{cases}$$

حيث أن $\propto, \beta$ تمثل معلمات التوزيع قاما، وأن $(\propto, \beta > 0)$

$\Gamma \propto$ تمثل دالة قاما ، وتأخذ دالة قاما الشكل: $\Gamma \propto = \int_0^\infty x^{\propto-1} e^{-\frac{x}{\beta}}$

وبصورة عامة، إذا كان n عدد صحيح موجب فإن دالة قاما تأخذ الشكل: $\Gamma n = (n-1)\,!$

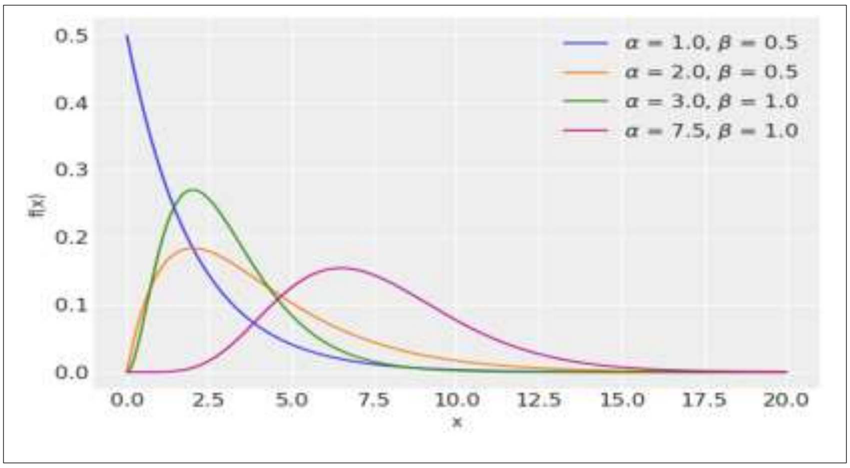
وفيما يلي بعض الحالات الخاصة للدالة قاما:

$$\Gamma_1 = 1$$

$$\Gamma_{\frac{1}{2}} = \pi$$

وغالبا ما يعبر عن توزيع قاما اختصارا بـ: $x \sim G(\propto, \beta)$

والشكل التالي يوضح دالة الكثافة الاحتمالية f(x) للتوزيع قاما عند بعض القيم لـ: $\propto, \beta$



ويمكن ايجاد دالة التوزيع التجميعية لتوزيع قاما كما يلي:

$$F(X) = P(X \leq x) = \int_0^x f(x) dx$$

$$F(X) = P(X \leq x) = \frac{1}{\beta^{\propto} \Gamma \propto} \int_0^x x^{\propto-1} e^{-\frac{x}{\beta}} dx$$

❖ **خصائص التوزيع قاما:**

- الوسط الحسابي: $\mu_x = \propto \beta$
- التباين: $\sigma_x^2 = \propto \beta^2$
- الانحراف المعياري: $\sigma_x = \beta\sqrt{\propto}$

مثال 05:

ليكن المتغير العشوائي x الذي يمثل مدة حياة آلة انتاجية (بالسنوات)، وله دالة التوزيع الاحتمالي التالية:

$$f(x) = \begin{cases} \frac{x}{4} e^{-\frac{x}{2}} & x \geq 0 \\ 0 & o/w \end{cases}$$

- اثبت أن دالة التوزيع الاحتمالي هي دالة كثافة احتمالية.
- ما هو احتمال ان تستمر الآلة في العمل مدة 10 سنوات أخرى على الاكثر؟
- احسب متوسط عمر الآلة وتباينه.

الحل:

- لإثبات أن دالة التوزيع الاحتمالي هي دالة كثافة احتمالية يجب تحقق الشروط التالية:

a) $f(x) \geq 0 \qquad \forall x \geq 0$

b) $\int_0^\infty f(x)dx = \int_0^\infty \frac{x}{4} e^{-\frac{x}{2}} dx = \frac{1}{4}\int_0^\infty x e^{-\frac{x}{2}} dx$

$$= \frac{1}{4}\left[-2e^{-\frac{x}{2}}\right]_0^\infty - \int_0^\infty \left(-2e^{-\frac{x}{2}}\right) dx$$

$$= \frac{1}{4}\left[0 + 2\int_0^\infty e^{-\frac{x}{2}} dx\right]$$

$$= 0 + \frac{1}{2}\left[-2e^{-\frac{x}{2}}\right]_0^\infty$$

$$= -[0 - 1]$$

$$\int_0^\infty f(x)dx = 1$$

ومنه دالة التوزيع الاحتمالي هي دالة كثافة احتمالية.

- حساب احتمال ان تستمر الآلة في العمل مدة 10 سنوات أخرى على الاكثر.

$$P(X \leq 10) = \int_0^{10} f(x)dx$$

$$= \frac{1}{4}\left[-2e^{-\frac{x}{2}}\right]_0^{10} + 2\int_0^{10} \left(e^{-\frac{x}{2}}\right) dx$$

$$= \frac{1}{4}\left\{[-20e^{-5}]_0^{10} + 2\left[-2e^{-\frac{x}{2}}\right]_0^{10}\right\}$$

$$= \frac{1}{4}\{(-0.14) - 4(0.007 - 1)\}$$

$$= \frac{1}{4}\{3.832\}$$

$$P(X \leq 10) = 0.958$$

- حساب متوسط عمر الآلة وتباينه: من خلال دالة التوزيع الاحتمالي نجد أن $(\propto = 2, \beta = 2)$

- الوسط الحسابي: $\mu_x = \propto \beta = 2(2) =$ 4 سنوات
- التباين: $\sigma_x^2 = \propto \beta^2 = 2(2)^2 =$ 8 سنوات مربع

4) التوزيع الأسي Exponential Distribution:

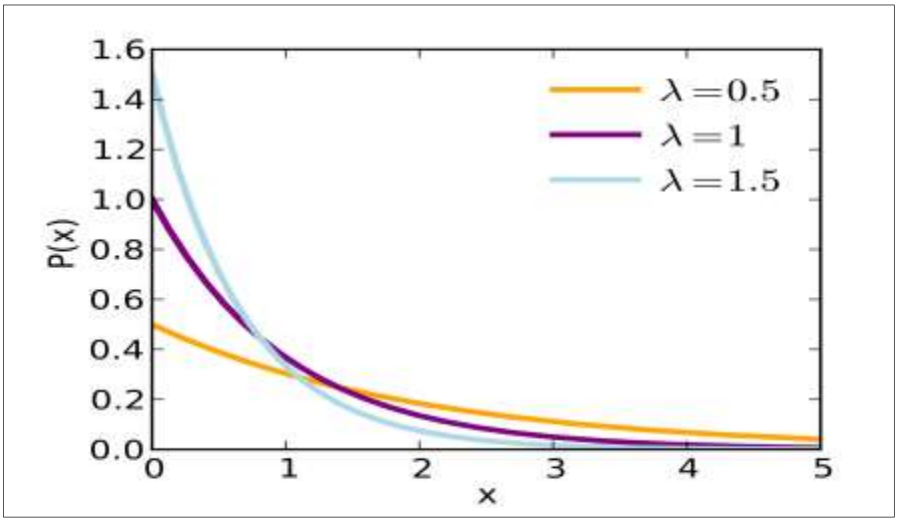
يعد التوزيع الأسي حالة خاصة من توزيع قاما وهذا عندما يكون $(\propto = 1)$، ويستخدم التوزيع الأسي لمعالجة بعض التطبيقات الاحصائية الخاصة بالمتغيرات العشوائية المتصلة، مثل مدة حياة بعض الاجزاء الالكترونية، فترة الانتظار عند الاشارة الضوئية، طبيعة البيانات المتعلقة بدرجات الحرارة القصوى والدنيا المسجلة من قبل الأرصاد الجوية...الخ.

وبافتراض أن لدينا متغير عشوائي x يتوزع وفق التوزيع الاسي، فإن دالة التويع الاحتمالي تاخذ الشكل:

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & o/w \end{cases}$$

وغالبا ما يعبر عن التوزيع الاسي اختصارا بالاصطلاح الآتي: $X \sim Exp(\beta)$

والشكل التالي يوضح دالة الكثافة الاحتمالية f(x) للتوزيع الاسي، حيث $\lambda = 1/\beta$:



ويمكن إيجاد دالة التوزيع التجميعية للتوزيع الاسي كما يلي:

$$F(X) = P(X \leq x) = \frac{1}{\beta}\int_0^x e^{\frac{-x}{\beta}} dx$$

$$= \frac{1}{\beta} -\beta e^{\frac{-x}{\beta}}\Big|_0^x$$

$$= -e^{\frac{-x}{\beta}}\Big|_0^x = 1 - e^{\frac{-x}{\beta}}$$

❖ خصائص التوزيع الأسي:

- الوسط الحسابي: $\mu_x = \beta$
- التباين: $\sigma_x^2 = \beta^2$
- الانحراف المعياري: $\sigma_x = \beta$

مثال 06:

إذا كانت مدة حياة ترانزيستور في جهاز التلفاز (بالساعة) عبارة عن متغير عشوائي له دالة كثافة احتمالية تأخذ الشكل التالي:

$$f(x) = \begin{cases} \frac{1}{500} e^{-\frac{x}{500}} & x \geq 0 \\ 0 & o/w \end{cases}$$

- ما هو احتمال أن يشتغل هذا الترانزيستور مدة 600 ساعة على الاكثر؟

- ما هو احتمال أن يشتغل ما بين 400 إلى 600 ساعة؟
- احسب متوسط عمر هذه القطعة، والتباين.

الحل:

$$P(X \leq x) = F(X)$$

$$P(X \leq 600) = F(600) = 1 - e^{\frac{-600}{500}} = 1 - 0.301$$

$$P(X \leq 600) = 0.699$$

$$P(400 \leq X \leq 600) = F(600) - F(400)$$

$$P(400 \leq X \leq 600) = \left[1 - e^{\frac{-600}{500}}\right] - \left[1 - e^{\frac{-400}{500}}\right]$$

$$= e^{\frac{-400}{500}} - e^{\frac{-600}{500}}$$

$$P(400 \leq X \leq 600) = 0.449 - 0.301 = 0.148$$

متوسط عمر القطعة الالكترونية هو: $\mu_x = \beta =$ 500ساعة

تباين عمر القطعة الالكترونية هو: $\sigma_x^2 = \beta^2 = 500^2 =$ 250000 ساعة مربع

5) توزيع بيتا Beta Distribution:

يعد توزيع بيتا من التوزيعات الاحتمالية المهمة في كثير من التطبيقات الاحصائية في الحياة العملية، ويستخدم على شكل واسع في دراسة سلوك بعض المتغيرات العشوائية للظواهر الطبيعية.

وبافتراض أن المتغير العشوائي المتصل x يتوزع وفق التوزيع بيتا، فإن دالة التوزيع الاحتمالي له تكون كما يلي:

$$f(x) = \begin{cases} \dfrac{\Gamma(\propto + \beta)}{\Gamma \propto \Gamma \beta} x^{\propto - 1}(1-x)^{\beta - 1} & 0 \leq x \leq 1 \\ 0 & o/w \end{cases}$$

وغالبا ما يعبر عن هذا التوزيع اختصارا بـ: $x \sim B(\propto, \beta)$

ويمكن إيجاد دالة التوزيع التجميعية لتوزيع بيتا كما يلي:

$$F(X) = P(X \leq x) = \int_0^1 f(x)dx$$

$$F(X) = \frac{\Gamma(\propto +\beta)}{\Gamma \propto \Gamma\beta} \int_0^1 x^{\propto-1}(1-x)^{\beta-1}dx$$

❖ خصائص التوزيع الاحتمالي بيتا:

- الوسط الحسابي: $\mu_x = \frac{\propto}{\propto+\beta}$
- التباين: $\sigma_x^2 = \frac{\propto\beta}{(\propto+\beta)^2(\propto+\beta+1)}$
- الانحراف المعياري: $\sigma_x = \frac{1}{\alpha+\beta} * \sqrt{\frac{\alpha\beta}{\propto+\beta+1}}$

مثال 07:

إذا كان المتغير العشوائي x الذي يمثل نسبة الرطوبة في مدينة المسيلة وله دالة التوزيع الاحتمالي التالية:

$$f(x) = \begin{cases} 30x^2(1-x)^2 & 0 \leq x \leq 1 \\ 0 & o/w \end{cases}$$

- أثبت أن دالة التوزيع الاحتمالي هي دالة كثافة احتمالية.
- ما هو احتمال أن تكون نسبة الرطوبة في مدينة المسيلة 40% على الأكثر؟
- ما هو احتمال أن تكون نسبة الرطوبة في مدينة المسيلة 30% على الأقل؟
- احسب متوسط نسبة الرطوبة والتباين لهذا التوزيع.

الحل:

- اثبات أن دالة التوزيع الاحتمالي هي دالة كثافة:

a) $f(x) \geq 0 \quad \forall x, \quad 0 \leq x \leq 1$

b) $\int_0^1 f(x)dx = \int_0^1 30x^2(1-x)^2dx = 30\int_0^1 x^2(1-2x+x^2)dx$

$$= 30\int_0^1 x^2 - 2x^3 + x^4)dx$$

$$= 30\left[\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\right]_0^1$$

$$= 30\left(\frac{1}{3} - \frac{1}{4} + \frac{1}{5}\right)$$

$$= 30\left(\frac{1}{3}\right) = 1$$

- حساب احتمال أن تكون نسبة الرطوبة في مدينة المسيلة 40% على الأكثر:

$$P(X \leq 0.4) = \int_0^{0.4} f(x)dx$$

$$= 30\int_0^{0.4}(x^2 - 2x^3 + x^4)dx$$

$$= 30\left[\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\right]_0^{0.4}$$

$$= 30\left(\frac{0.4^3}{3} - \frac{0.4^4}{4} + \frac{0.4^5}{5}\right)$$

$$= 30(0.021 - 0.013 + 0.002)$$

$$P(X \leq 0.4) = 0.3$$

- حساب احتمال أن تكون نسبة الرطوبة في مدينة المسيلة 30% على الأقل:

$$P(X \geq 0{,}3) = 1 - \int_0^{0.3} f(x)dx$$

$$= 1 - 30\int_0^{0.3}(x^2 - 2x^3 + x^4)dx$$

$$= 1 - 30\left[\frac{x^3}{3} - \frac{x^4}{2} + \frac{x^5}{5}\right]_0^{0.3}$$

$$= 1 - 30\left(\frac{0.3^3}{3} - \frac{0.3^4}{4} + \frac{0.3^5}{5}\right)$$

$$= 1 - 30(0{,}0054)$$

$$P(X \geq 0{,}3) = 0.838$$

- حساب متوسط نسبة الرطوبة والتباين لهذا التوزيع.

▪ الوسط الحسابي: $\mu_x = \frac{\propto}{\propto+\beta} = \frac{3}{3+3} = 0.5$

أي أن متوسط الرطوبة في المدينة هو 50%.

▪ التباين: $\sigma_x^2 = \frac{\propto\beta}{(\propto+\beta)^2(\propto+\beta+1)} = \frac{3*3}{(3+3)^2+(3+3+1)} = 0.036$

6) توزيع كاي مربع χ^2:

يعتبر توزيع كاي تربيع χ^2 من التوزيعات الاحتمالية الشائعة الاستخدام حيث توجد له تطبيقات عديدة بدرجة يمكن معها القول أنه يأتي في المرتبة الثانية للتوزيع المعتدل من حيث كثرة تطبيقاته، ونذكر على سبيل المثال اختبارات جودة المطابقة والتجانس والاستقلالية والتباين ...الخ.

ويعتمد توزيع χ^2 مثل توزيع t اعتمادا كاملا على درجات الحرية،

فكلما زادت درجات الحرية كلما قل التواء التوزيع واقترب من التماثل.

❖ دالة الكثافة الاحتمالية:

إذا كان المتغير العشوائي x يتبع توزيع كاي تربيع $X \sim \chi_n^2$, فإن دالة الكثافة الاحتمالية ل x لما $x > 0$ تعطى بالشكل:

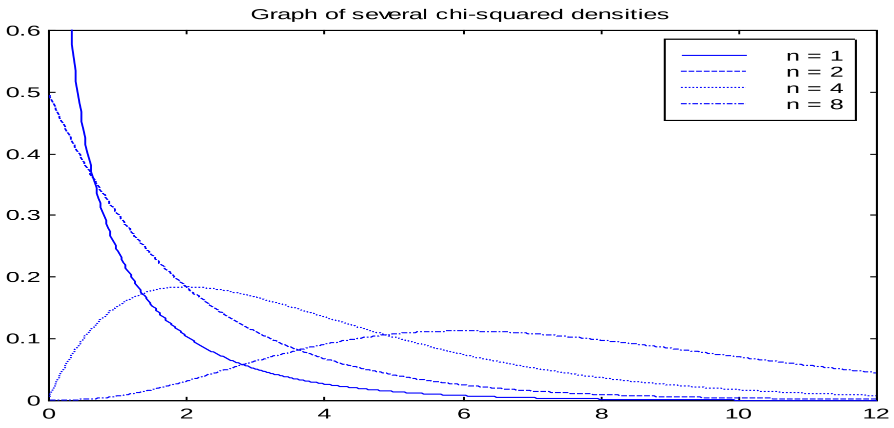
$$.f(x)=\begin{cases} f(x)=\dfrac{1}{2^{n/2}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-x/2} & x\geq 0 \\ 0 & x\leq 0 \end{cases}$$

حيث:

$$\Gamma(\propto) = \int_0^{\infty} \chi^{\propto-1}e^{-\chi}d\chi \qquad \propto > 0$$

وتحدد درجة الحرية n شكل التوزيع والتوائه، إذا كانت n أقل من 3 فإن منوال دالة الكثافة يكون في الصفر 0، وإذا كانت n أكبر من 3 فإن المنوال ينزاح بعيدا عن الصفر 0 كلما زادت قيمة n، والشكل الموالي يبين بعض المنحنيات لتوزيع كاي مربع لدرجات حرية مختلفة.

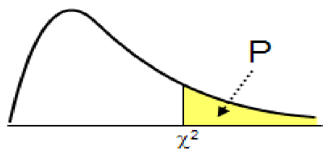
شكل توزيع كاي تربيع χ^2



❖ خصائص توزيع كاي مربع:

- المتوسط الحسابي (التوقع الرياضي): $E(x) = \mu = n$
- التباين: $V(x) = \sigma^2 = 2n$

جدول توزيع كاي تربيع χ^2



	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588
11	2.603	3.816	14.631	17.275	19.675	21.920	22.618	24.725	26.757	29.354	31.264
12	3.074	4.404	15.812	18.549	21.026	23.337	24.054	26.217	28.300	30.957	32.909
13	3.565	5.009	16.985	19.812	22.362	24.736	25.472	27.688	29.819	32.535	34.528
14	4.075	5.629	18.151	21.064	23.685	26.119	26.873	29.141	31.319	34.091	36.123
15	4.601	6.262	19.311	22.307	24.996	27.488	28.259	30.578	32.801	35.628	37.697
16	5.142	6.908	20.465	23.542	26.296	28.845	29.633	32.000	34.267	37.146	39.252
17	5.697	7.564	21.615	24.769	27.587	30.191	30.995	33.409	35.718	38.648	40.790
18	6.265	8.231	22.760	25.989	28.869	31.526	32.346	34.805	37.156	40.136	42.312
19	6.844	8.907	23.900	27.204	30.144	32.852	33.687	36.191	38.582	41.610	43.820
20	7.434	9.591	25.038	28.412	31.410	34.170	35.020	37.566	39.997	43.072	45.315
21	8.034	10.283	26.171	29.615	32.671	35.479	36.343	38.932	41.401	44.522	46.797
22	8.643	10.982	27.301	30.813	33.924	36.781	37.659	40.289	42.796	45.962	48.268
23	9.260	11.689	28.429	32.007	35.172	38.076	38.968	41.638	44.181	47.391	49.728
24	9.886	12.401	29.553	33.196	36.415	39.364	40.270	42.980	45.559	48.812	51.179
25	10.520	13.120	30.675	34.382	37.652	40.646	41.566	44.314	46.928	50.223	52.620
26	11.160	13.844	31.795	35.563	38.885	41.923	42.856	45.642	48.290	51.627	54.052
27	11.808	14.573	32.912	36.741	40.113	43.195	44.140	46.963	49.645	53.023	55.476
28	12.461	15.308	34.027	37.916	41.337	44.461	45.419	48.278	50.993	54.411	56.892
29	13.121	16.047	35.139	39.087	42.557	45.722	46.693	49.588	52.336	55.792	58.301
30	13.787	16.791	36.250	40.256	43.773	46.979	47.962	50.892	53.672	57.167	59.703
31	14.458	17.539	37.359	41.422	44.985	48.232	49.226	52.191	55.003	58.536	61.098

❖ حساب الاحتمالات من توزيع كاي مربع:

يمكن استخراج الاحتمالات المرتبطة بتوزيع كاي تربيع ذو درجة حرية معينة من الجداول الإحصائية، حيث تتركز درجات الحرية بالتدرج على العمود الأيسر للجدول، وتبعا لكل درجة حرية يمكن استخراج قيمة χ^2 الموافقة لاحتمال معين، أين تتموقع الاحتمالات في السطر العلوي للجدول.

مثال 08:

- إذا كان لدينا توزيع كاي تربيع ذو درجة حرية 3 فإن:

$X \sim \chi^2_3$ فإن $P(X < 2.5) = 0.5247$

- إذا كان $X \sim \chi^2_7$ فأوجد: $P(X > 10)$

من الجدول الاحصائي نجد:

$$P(X < 10) = 0.8114 \Rightarrow P(X > 10) = 1 - 0.8114 = 0.1886$$

❖ النقطة الحرجة:

يتم الاشارة الى النقطة $100\alpha\ \%$ للتوزيع χ^2_n ذو درجة الحرية n أي t_n والتي يرمز لها بالرمز $\chi^2_{n,\alpha}$، فإذا كان $X \sim \chi^2_n$ فإن:

$$P(X > \chi^2_{n,\alpha}) = \alpha.$$

يتم جدولة النقطة الحرجة للتوزيع χ^2 بشكل منفصل، بتلاقي السطر الذي يبدأ بدرجة الحرية المعينة، مع العمود الذي يحوي المساحة $100\alpha\ \%$

ويعتبر توزيع χ^2 توزيعا غير متناظر مثل التوزيع الطبيعي وتوزيع ستيودنت، وبالتالي إذا أردنا إيجاد النقطة الحرجة للمساحة الأدنى من قيمة معينة (مثلا إيجاد قيمة x لما $P(X < x) = \alpha$)، وبالتالي لا نستطيع استخراجها مباشرة من النقطة الحرجة العليا، ولكننا نحتاج إلى إيجاد $\chi^2_{n,1-\alpha}$

مثال 09:

- إذا كان $\chi^2_{n,1-\alpha}$ أوجد النقطة الحرجة لأدنى 01% (أي قيمة x حيث $P(X < x) = 0.01$)

النقطة الحرجة لأدنى 01% يرمز لها بـ $\chi^2_{8,0.99}$ والتي تقابلها القيمة 1.646.

- نفرض أن $X \sim \chi^2_{10}$ أوجد قيمة x حيث $P(X > x) = 0.1321$

هنا النقطة x يجب أن يكون فوقها 13.21% من مساحة المنحني، لكن 0.1321 ليست قيمة معيارية للمساحة α ومنه فنحن بحاجة إلى استعمال جدول دالة التوزيع الاحتمالي فإيجاد x.

$P(X > x) = 0.1321 \Rightarrow P(X < x) = 1 - 0.1321 = 0.8679$

ومن جدول التوزيع الاحتمالي نجد أن قيمة x هي: x=15

7) توزيع ستودنت t:

تعريف: بافتراض أن لدينا متغيرين عشوائيين مستقلين Y، Z حيث:

$$Y \sim N[0, 1] \quad \text{و} \quad Z \sim \chi^2_n .$$

فإن المتغير العشوائي x المعرف بـ:

$$X = \frac{Y}{\sqrt{Z/n}}$$

يتبع توزيع ستيودنت ذو n درجة حرية ويرمز له ب: t_n ودالة الكثافة التالية:

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \qquad -\infty < t < \infty$$

حيث:

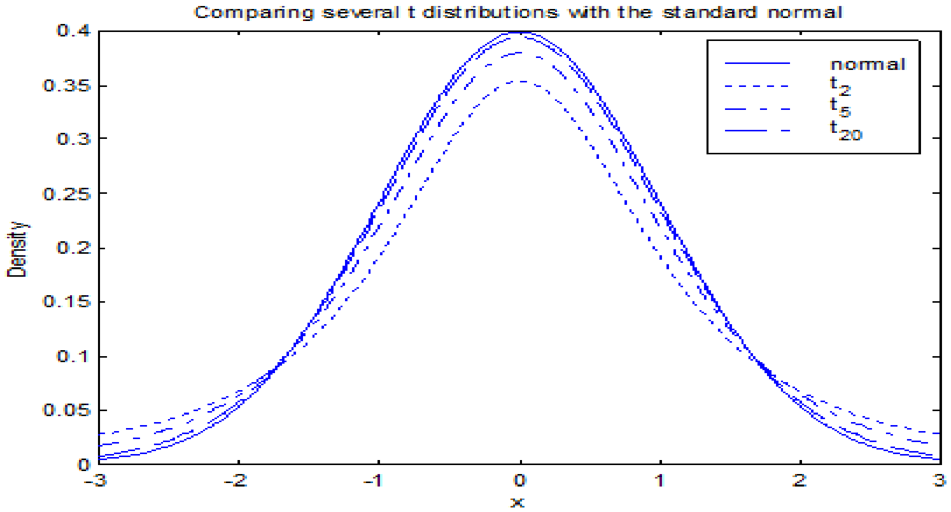
$$\Gamma(\propto) = \int_0^\infty t^{\propto-1}e^{-t}dt \qquad \propto > 0$$

يعتبر توزيع t متماثل حول الصفر وشكله العام يشبه شكل الجرس للتوزيع الطبيعي. ومع ذلك، فإن ذيول توزيع t يمكن أن تقترب من الصفر بشكل أبطأ بكثير من التوزيع الطبيعي – أي أن توزيع t هو ذيل ثقيل أكثر من المعتاد. تحدد درجات الحرية مدى ثقل الذيل في توزيع t.

ملاحظة: يُشار أحيانًا إلى توزيع t مع n = 1 باسم توزيع كوشي. هذا ذيل ثقيل لدرجة أن متوسطه وتباينه غير موجودين، (هذا لأن التكاملات التي تحدد المتوسط والتباين ليست متقاربة تمامًا.)

ملاحظة مهمة: تتقارب كثافة توزيع ستيودنت مع كثافة التوزيع الطبيعي لما تؤول n إلى مالا نهاية $n \to \infty$،
وعموما يعتبر الاحصائيون أن المنحنيان يتطابقان تقريبا عند $n \geq 30$

يوضح الرسم البياني أدناه كيف يختلف توزيع t باختلاف درجات الحرية.



يمكن البحث عن الاحتمالات المرتبطة بتوزيع t في الجداول الإحصائية، حيث يتم الإشارة إلى درجات الحرية مرة أخرى بواسطة v ويتم سردها بالتدرج في العمود الأيسر من الجدول، بينما تأتي المساحات التي على يمين القيمة t في السطر العلوي من الجدول، ثم لكل قيمة t ذات درجة حرية معينة مساحة مقابلة لها في السطر العلوي من الجدول.

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

مثال 10:

اذا كان لدينا $X \sim t_3$ فإن $P(X < 2.5) = 0.9561$

مثال 11:

إذا كان $X \sim t_{12}$ أوجد $P(X > 2.5)$

- $P(X > 2.5) = 1 - P(X < 2.5) = 1 - 0.986 = 0.014$

❖ خصائص التوزيع t:

- المتوسط: $\mu = 0$
- التباين: $\sigma^2 = \frac{n}{n-2} \quad n > 2$
- الانحراف المعياري: $\sigma = \sqrt{\frac{n}{n-2}} \quad n > 2$

❖ النقطة الحرجة:

يتم الاشارة الى النقطة % 100α للتوزيع t ذو درجة الحرية n أي t_n والتي يرمز لها بالرمز $t_{n,\alpha}$ ، فإذا كان $X \sim t_n$ فإن:

$$P(X > t_{n,\alpha}) = \alpha.$$

يتم جدولة النقطة الحرجة للتوزيع t بشكل منفصل، بتلاقي السطر الذي يبدأ بدرجة الحرية المعينة، مع العمود الذي يحوي المساحة % 100α

مثال 12:

- أوجد النقطة الحرجة %05 لما يكون التوزيع t_6.

مباشرة من الجدول يمكننا النظر إلى أن: $t_{6,0.05} = 1.943$. وهذا يعني أن :

$$P(X > 1.943) = 0.05$$

ونظرا لأن توزيع t متماثل، فإن العثور على النقطة الحرجة الأقل أمر بسيط.

مثال 13:

لدينا: $X \sim t_{10}$. أوجد قيمة t حيث $P(X < t) = 0.01$ (أي بعبارة أخرى أوجد النقطة التي تحتها أقل %01).

النقطة التي فوقها أعلى %01 هي $t_{10,0.01} = 2.764$. وبما أن:

$$P(X > 2.764) = 0.01 \Rightarrow P(X < -2.764) = 0.01$$

إذن النقطة t التي تحتها أقل %01 هي: -2.764

8) توزيع فيشر F:

يعد توزيع F من التوزيعات الاحصائية الهامة حيث يستخدم في الاحصاء الاستنتاجي (الاستدلالي)، لاجراء العديد من اختبارات الفروض المتعلقة بتحليل التباين وتصميم التجارب واختبار معنوية معلمات الانحدار،

وقد تم اكتشافه من قبل الاحصائي الشهير فيشر R.A. Fisher، حيث استخدمه لاختبار النسبة بين تبايني مجتمعين طبيعيين.

تعريف: بفرض أنه لينا متغيرين عشوائيين z، y بحيث أنه: $nY \sim \chi_n^2$ و $mZ \sim \chi_m^2$.

المتغير العشوائي x المعرف بما يلي: $X = \frac{Y}{Z}$ نقول عنه أنه يتبع توزيع فيشر ذو درجتي الحرية n ، m

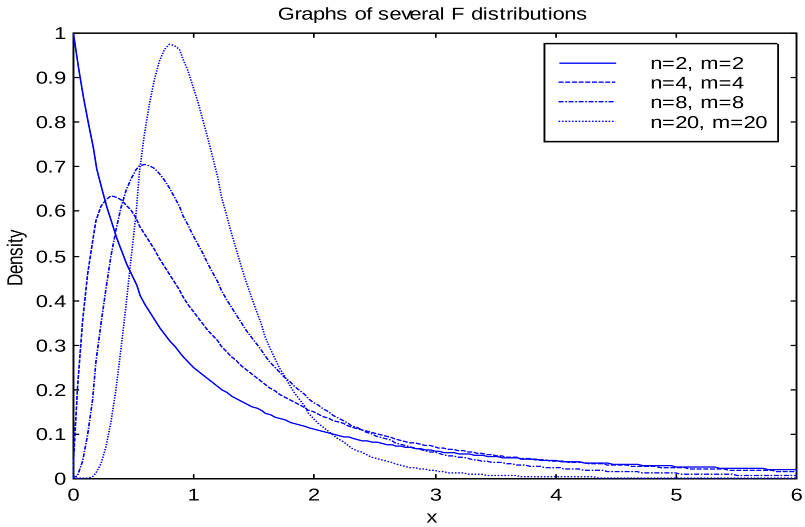
ونرمز له بالرمز $F_{n,m}$

ودالة الكثافة للتوزيع فيشر هي على الشكل:

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} n^{\frac{n}{2}} m^{\frac{m}{2}} x^{\left(\frac{n}{2}-1\right)} (n+mx)^{-\frac{n+m}{2}} & x > 0 \\ 0 & x < 0 \end{cases}$$

لذلك فإن التوزيع F له معلمتان وهما درجتي حرية، وترتيب درجتي الحرية مهم جدا، فالتوزيع $F_{n,m}$ ليس هو نفسه التوزيع $F_{m,n}$

ملاحظة: يتم تحديد دالة كثافة التوزيع F فقط للقيم الموجبة لـ x، وتحدد قيم درجتي الحرية شكل التوزيع. والشكل التالي يعرض منحنيات توزيع F لقيم مختلفة لـ n و m أدناه:



وكلما زادت درجات حرية التوزيع F اقترب شكله من شكل التوزيع الطبيعي، وتستخدم الجداول الاحصائية للتوزيع F في الملاحق لحساب المساحات (الاحتمالات) تحت منحنى التوزيع، والرمز $F_{(\gamma;\, v_1;\, v_2)}$ يعبر عن النقطة على المحور الافقي التي يكون على يسارها المساحة γ .

❖ خصائص التوزيع F:

- المتوسط: $m > 2$ $\quad \mu = \frac{m}{m-2}$
- التباين: $m > 4$ $\quad \sigma^2 = \frac{2m^2(n+m-2)}{n(m-4)(m-2)^2}$
- الانحراف المعياري: $\sigma = \sqrt{\frac{2m^2(n+m-2)}{n(m-4)(m-2)^2}}$

❖ النقطة الحرجة:

يمكن العثور على جداول منفصلة تعطي 10 و 5 و 2.5 و 1 و 0.5 و 0.1 نقطة مئوية لتوزيعات F مع مجموعات مختلفة من درجات الحرية.

سوف نشير الى النقطة (التي فوقها المساحة % 100α) للتوزيع $F_{n,m}$ بالرمز $F_{n,m,\alpha}$، لما يكون $X \sim F_{n,m}$، إذن $P(X > F_{n,m,\alpha}) = \alpha$.

في الجدول الإحصائي لـ: % 100α للتوزيع F، يتم الاشارة لدرجة الحرية الأولى (درجة حرية البسط)، ويتم سردها في السطر العلوي للجدول، بينما تتم الاشارة لدرجة الحرية الثانية (درجة حرية المقام)، ويتم سردها في العمود الايسر للجدول.

مثلا إيجاد النقطة الحرجة 1% للتوزيع F ذو درجتي الحرية 5 و 3.

لدينا الجدول الاحصائي 1% لتوزيع فيشر، ومنه نستخرج القيمة وهي: $F_{5,3,0.01} = 28.24$.

$v_1 =$		1	2	3	4	5
$v_2 =$	**1**	4052	4999	5403	5625	5764
	2	98.50	99.00	99.17	99.25	99.30
	3	34.12	30.82	29.46	28.71	**28.24**
	4	21.20	18.00	16.69	15.98	15.52
	5	16.26	13.27	12.06	11.39	10.97

مثال 14:

أوجد النقطة الحرجة 5% في كلا التوزيعين، $F_{5,10}$ و $F_{10,5}$

من الجدول 5% نجد $F_{5,10,0.05} = 3.326$ و $F_{10,5,0.05} = 4.735$

ملاحظة: إن الجداول الاحصائية تعطينا النقطة الحرجة التي أعلاها مساحة معينة فقط، أي:

$P(X > x) = \alpha$، من أجل قيم صغيرة لـ α، ونظرا لأن توزيع فيشر غير متناظر فإنه ولإيجاد نقطة مئوية تحتها مساحة معينة، لا يمكننا استخدام سالب النقطة الحرجة المقابلة مثل توزيعي ستيودنت والتوزيع الطبيعي.

$$P(X > x) \neq P(X < -x).$$

نتيجة: بفرض أن $X = \frac{Y}{Z} \sim F_{n,m}$ فإنه يصبح $X^{-1} = \frac{Z}{Y} \sim F_{m,n}$.

وهذا انطلاقا من تعريف التوزيع F، حيث أنه إذا كان $nY \sim \chi_n^2$ and $mZ \sim \chi_m^2$. فإن $X = \frac{Y}{Z} \sim F_{n,m}$

ومن تعريف التوزيع F فإن $\frac{Z}{Y} \sim F_{m,n}$ كما هو مطلوب، أي $X^{-1} = \frac{Z}{Y} \sim F_{m,n}$.

ويمكننا استخدام هذه النتيجة لإيجاد نقاط مئوية أقل (على يسارها مساحة معينة) لتوزيعات F.

نتيجة مهمة: إن النقطة الحرجة التي على يسارها مساحة معينة α100% للتوزيع $F_{n,m}$، هي مقلوب النقطة الحرجة التي على يمينها مساحة معينة α100% للتوزيع $F_{m,n}$

مثال 15:

- ليكن لدينا: $X \sim F_{5,10}$. ونريد ايجاد قيمة النقطة x حيث $P(X < x) = 0.05$، أي إيجاد النقطة التي على يسارها المساحة 5%.

إذن النقطة الحرجة التي على يسارها المساحة 5% للتوزيع $F_{5,10}$، هي مقلوب النقطة الحرجة التي على يمينها المساحة 5% للتوزيع $F_{10,5}$

إذن $$x = \frac{1}{F_{10,5,0.05}} = \frac{1}{4.735} = 0.2112$$

- نفرض أن $X \sim F_{4,7}$، أوجد النقطتين التي على يمينهما وعلى يسارهما 10% من المساحة.

النقطة التي على يمينها (أعلاها 10% من المساحة) يمكن استخراجها مباشرة من الجدول 10%.

$$F_{4,7,0.1} = 2.961.$$

النقطة التي على يسارها (أدناها 10% من المساحة) قيمتها هي مقلوب قيمة النقطة التي أعلاها 10% من المساحة (ولكن في التوزيع $F_{7,4}$) أي:

$$F_{4,7,0.9} = \frac{1}{F_{7,4,0.1}} = \frac{1}{3.979} = 0.2513$$

مجموعة من التمارين المحلولة:

التمرين 01:

إذا كانت المتغيرة x تتبع التوزيع الطبيعي المعياري، فما هي احتمالات قيم z في الحالات التالية:

- احتمال الحصول على قيمة x تقع بين 0 و 0.5.
- احتمال الحصول على قيمة x تقع بين −0.5 و 1.5.
- احتمال الحصول على قيمة x تقع بين 1 و 2.5.
- احتمال الحصول على قيمة x تقع بين −0.7 و 0.7.
- ما قيمة x التي احتمال الحصول على قيمة أعلى منها هو 0.778.
- ما قيمتي x اللتان تحصران 90% من المساحة تحت المنحني.

الحل:

- $P(0 < z < 0.5) = \emptyset(0,5) - \emptyset(0)$

$= 0.6915 - 0.50$

$P(0 < z < 0.5) = 0.1915$

- $P(-0.5 < z < 1.5) = \emptyset(1,5) - \emptyset(-0.5)$

$= 0.9332 - 0.3085$

$P(-0.5 < z < 1.5) = 0.6247$

- $P(1 < z < 2.5) = \emptyset(2,5) - \emptyset(1)$

$= 0.9938 - 0.8413$

$P(1 < z < 2.5) = 0.1525$

- $P(-0.7 < z < 0.7) = \emptyset(0.7) - \emptyset(-0.7)$

$= 0.7580 - 0.2420$

$P(-0.7 < z < 0.7) = 0.5160$

- قيمة z_0 التي احتمال الحصول على قيمة اعلى منها هو 0.778 هي: $P(z>z_0)=0.778$;

$z_0=-0.76$

- قيمتي z اللتان تحصران 90% من المساحة تحت المنحني الطبيعي المعياري هما:

قيمتي z اللتان أكبر منهما وأقل منهما مساحة 10% مقسومة على اثنين أي 05% في كل جهة، وبالتالي هما -1.64 و$+1.64$.

التمرين 02:

تتوزع صفة الوزن لأطفال جمعية ذات ال5 سنوات، توزيعا طبيعيا بمتوسط قدره 18 كغ، وتباين قدره 4 كغ^2، والمطلوب:

- ما هو احتمال الحصول على طفل وزنه بين 16–20 كغ، 14–18 كغ، 18–24 كغ.

وما هي أوزان الاطفال التي تحصر 95% من الأفراد.

الحل:

لدينا: $x \sim N(18,2)\ ; \quad \mu = 18\ ; \quad \sigma^2 = 4\ ; \quad \sigma = 2$

1- $P(16 \le X \le 20) = P\left(\frac{16-18}{2} \le \frac{X-\mu}{\sigma} \le \frac{20-18}{2}\right)$

$$P(16 \le X \le 20) = P(-1 \le z \le 1)$$
$$= \phi(1) - \phi(-1)$$
$$= 0.8413 - 0.1587$$
$$= 0.6826$$

2- $P(14 \le X \le 18) = P\left(\frac{14-18}{2} \le \frac{X-\mu}{\sigma} \le \frac{18-18}{2}\right)$

$$P(14 \le X \le 18) = P(-2 \le z \le 0)$$
$$= \phi(0) - \phi(-2)$$
$$= 0.5 - 0.0228$$
$$= 0.4772$$

3- $P(18 \le X \le 24) = P\left(\frac{18-18}{2} \le \frac{X-\mu}{\sigma} \le \frac{24-18}{2}\right)$

$$P(18 \le X \le 24) = P(0 \le z \le 3)$$
$$= \phi(3) - \phi(0)$$
$$= 0.9987 - 0.5$$
$$= 0.4987$$

4- أوزان الأطفال التي حصر 95% من الأطفال هي:

$$P(x_1 \le X \le x_2) = P\left(\frac{x_1 - 18}{2} \le \frac{X-\mu}{\sigma} \le \frac{x_2 - 18}{2}\right) = 0.95$$

$$\Rightarrow \phi(z_2) - \phi(z_1) = 0.95$$

$$\Rightarrow \phi(z_1) = 0.05/2 = 0.025$$

$$\Rightarrow z_1 = -1.96$$

$$\Rightarrow z_2 = 1.96$$

$$\frac{x_1 - 18}{2} = -1.96 \Rightarrow x_1 = (-1.96 \times 2) + 18 = 14.08$$

$$\frac{x_2 - 18}{2} = 1.96 \Rightarrow x_2 = (1.96 \times 2) + 18 = 21.92$$

أي أن أوزان الاطفال التي تحصر 95% من الأطفال تكون بين 14,08 كغ و21,92 كغ.

<u>التمرين 03:</u>

أ) أوجد قيم t عند درجة حرية 18 التي تحقق الاحتمالات التالية:

- المساحة المظللة عن اليمين هي 0.10.
- المساحة المظللة عن اليسار هي 0.99.

ب) أوجد القيم الحرجة لـ t والتي تجعل المساحة في الطرف الأيمن لتوزيع t هي 0.05، إذا كانت درجة الحرية تساوي: 16، 27، 120.

<u>الحل:</u>

أ) في توزيع t ذو درجة الحرية 18، تكون النقطة t_0 التي عن يمينها 0.10 أو 10% من المساحة تحت المنحني t، هي $t_0=1.33$

والنقطة t_1 التي عن يسارها 0.99 أو 99% من المساحة تحت المنحني t، هي نفسها النقطة التي عن يمينها 0,01، أي $t_1=2.552$.

ب) $t_{(0.05\ ;\ 16)}= 1.746$

$t_{(0.05\ ;\ 27)}= 1.703$

$t_{(0.05\ ;\ 120)}= 1.658$

<u>التمرين 04:</u>

في مصنع لإنتاج المصابيح تبين أن عدد ساعات عمل المصابيح تتوزع طبيعيا، بمتوسط قدره 500 ساعة، وانحراف معياري قدره 10 ساعات.

المطلوب هو حساب:

- احتمال أن يعمل مصباح بين 500 ساعة و515 ساعة.
- احتمال أن يعمل مصباح أكبر من 520 ساعة.
- احتمال أن يعمل مصباح أقل من 480 ساعة.

<u>الحل:</u>

نفرض أن المتغير العشوائي x يعبر عن العمر الافتراضي للمصابيح المنتجة، أي أن:

المتوسط μ = 500 ساعة

الانحراف المعياري σ = 10 ساعة

1− احتمال أن يكون العمر الافتراضي للمصباح بين 500 ، 515 ساعة

$$P(500 \leq X \leq 515) = P\left(\frac{500-500}{10} \leq \frac{X-\mu}{\sigma} \leq \frac{515-500}{10}\right)$$

$$= \mathrm{P}(0 \leq \mathrm{Z} \leq 1.5) = 0.4332$$

2− احتمال أن يكون العمر الافتراضي للمصباح أكبر من 520 ساعة

$$P(X > 520) = P\left(\frac{X-\mu}{\sigma} > \frac{520-500}{10}\right) = P(Z > 2)$$

$$= P(Z > 2) = P(Z < -2)$$

$$P(x > 510) = 0.0228$$

3− احتمال أن يكون العمر الافتراضي للمصباح أقل من 480 ساعة

$$P(X < 480) = P\left(\frac{X-\mu}{\sigma} < \frac{480-500}{10}\right)$$

$$= P(z < -2)$$

$$= \phi(-2)$$

$$P(X < 480) = 0.0228$$